Chapter 9

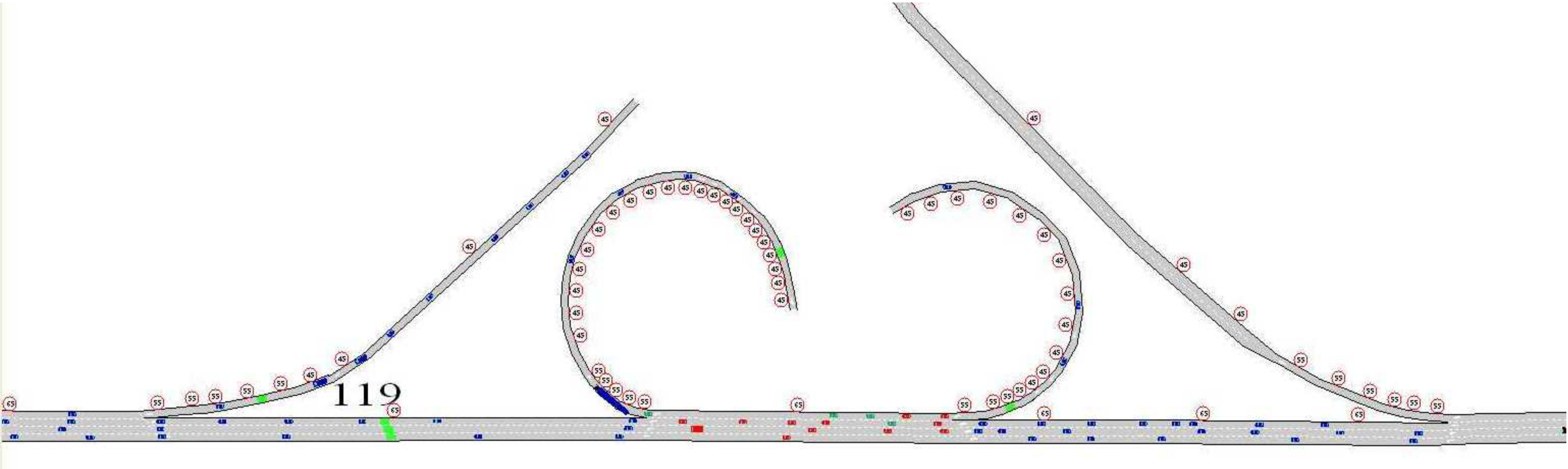
Model Calibration

John Hourdakis
Center for Transportation Studies, U of Mn
Hourd001@tc.umn.edu

Why Calibrate?

- Computers Cannot Magically Replicate Reality!
  - Simulation Models Are Designed to be General
  - Driver Behavior and Road Characteristics Depend on
    - Location i.e. Minnesota vs California
    - Vehicle Characteristics (Horsepower, Size, etc.)
    - Weather Conditions (Dry, Wet, Ice, etc.)
    - Day or Night
  - Microscopic Simulators Can Adapt and Replicate Almost Any Condition if the Model Parameters Are Properly Adjusted
  - What is Realistic and What is Not?

Example of a Disaster



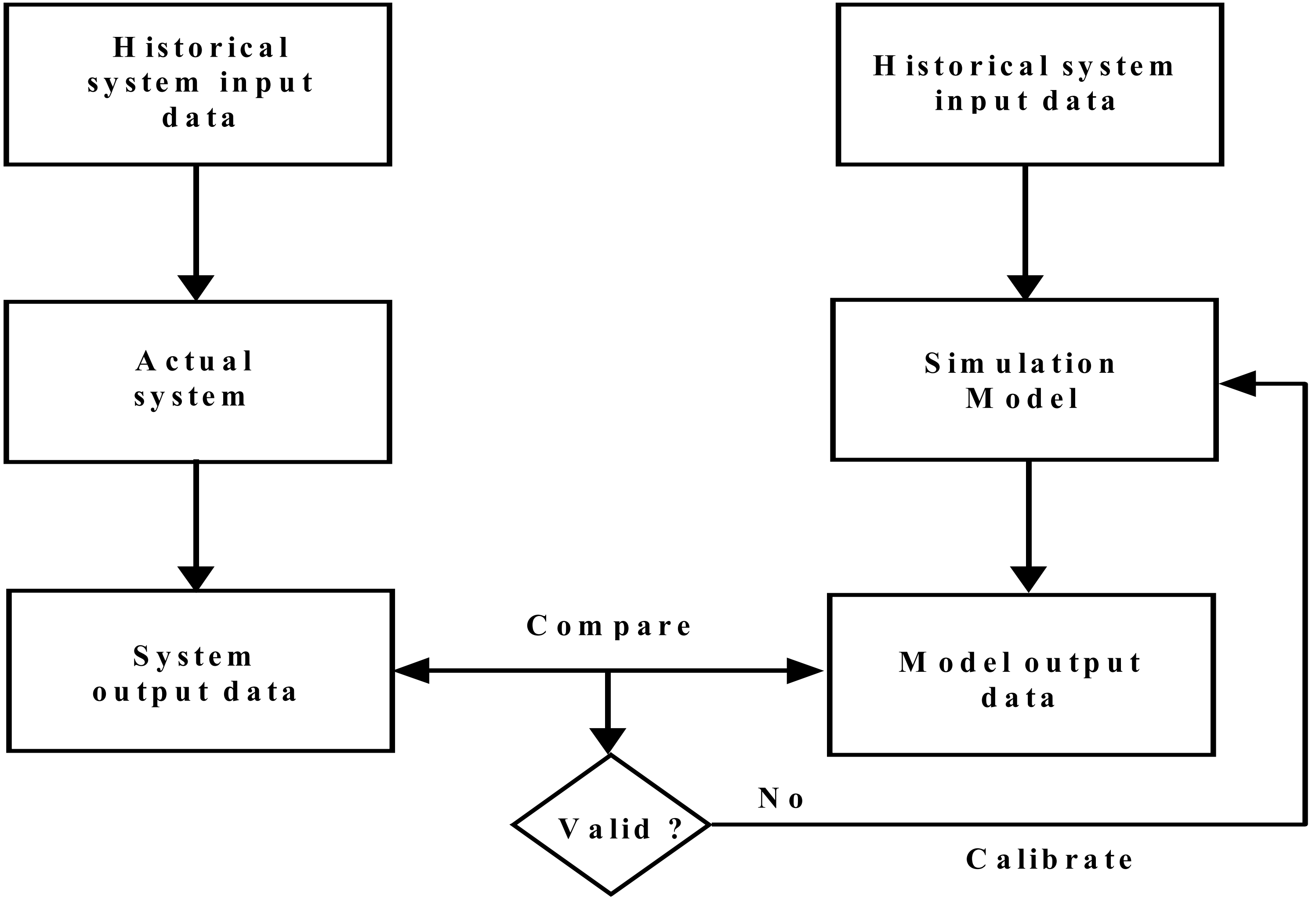
- Half of an Ordinary Double Cloverleaf
- Freeway Speed 65 mph, Ramps 45 mph
- Turning Warnings More Than 2000 Feet Away
- Lot’s of Traffic: 6500 veh/hour on Mainline
- Three Vehicle Types: Car, Truck, and Semi-Trailer

Simulation Results From 5 Scenarios

- Same Demands, Speed Limits, Turnings, Most of Model Parameters.
- Changed Car and Semi-Trailer :
  - Acceleration
  - Normal Deceleration
  - Maximum Deceleration (Emergency Stop)

	Average Speed	Average Flow	Total Travel Time	Av. Delay	LOS
$+2f/s^2$	45.5 mph	6831 vph	257 hours	9 sec/veh	D
$+1f/s^2$	29.1 mph	6200 vph	484 hours	41 sec/veh	F
BASE	46.7 mph	6396 vph	234 hours	9 sec/veh	D
$-1f/s^2$	35.3 mph	6152 vph	393 hours	30 sec/veh	E
$-2f/s^2$	27.4 mph	5500 vph	536 hours	57 sec/veh	F

Calibration Procedure



REALITY
Current Traffic Measurements, etc.

- Need to KNOW the Network to be Replicated
- Depending on the Scope of the Project, Two or More of the Following Are Needed :
  - Mainline Volumes (Every X Feet)
  - Mainline Speeds (Every X Feet)
  - Travel Times (Link or Between O/D Pairs)
  - Bottleneck Capacity (Measured, Not Theoretical)
  - Entrance Ramp Queues
  - Intersection Queues and Queue Discharge Rates

Modeling Parameters

- **Global Parameters**
  - Vehicle Parameters
    - Size and Power
    - Driver Behavior
  - Car-Following Model Parameters
    - Reaction Time
    - Speed Distibution Among Lanes (Overtaking Manuevers)
- **Local Section Parameters**
  - Curvature and Grade
  - Speed Limit or Free Flow Speed
  - Lane Changing Distances
  - Headway/Hesitation Factor

Model Parameter Issues

- Driver/Vehicle Characteristics
  - Literature Does Not Provide Adequate Information
  - Not Common to All Simulators
  - No Info on Effects of Weather/Pavement/Ambient Conditions
- Car-Following Parameters
  - Depend on the Simulator’s Car-Following Model Employed.
  - Most Simulators Do Not Adequately Describe the Modeling Process (if at all).
    - User Has No Sense of Model Parameter Effects on Results.

Calibration Issues

- Very Important For Model Accuracy and Robustness
- Accuracy Depends on Measurement Granularity
  - Averages Over Several Days is a Bad Choice
  - Might Need Additional Information to be Collected in Turbulent Sections (Bottlenecks, Weaving Areas, etc.)
- Simulation Objective Affects Calibration
  - When Adaptive Control Strategies Are Simulated, Stricter Validation is Needed
  - Modeling of an Isolated Interchange in Rural Minnesota Will be Restrictive

Calibration Issues

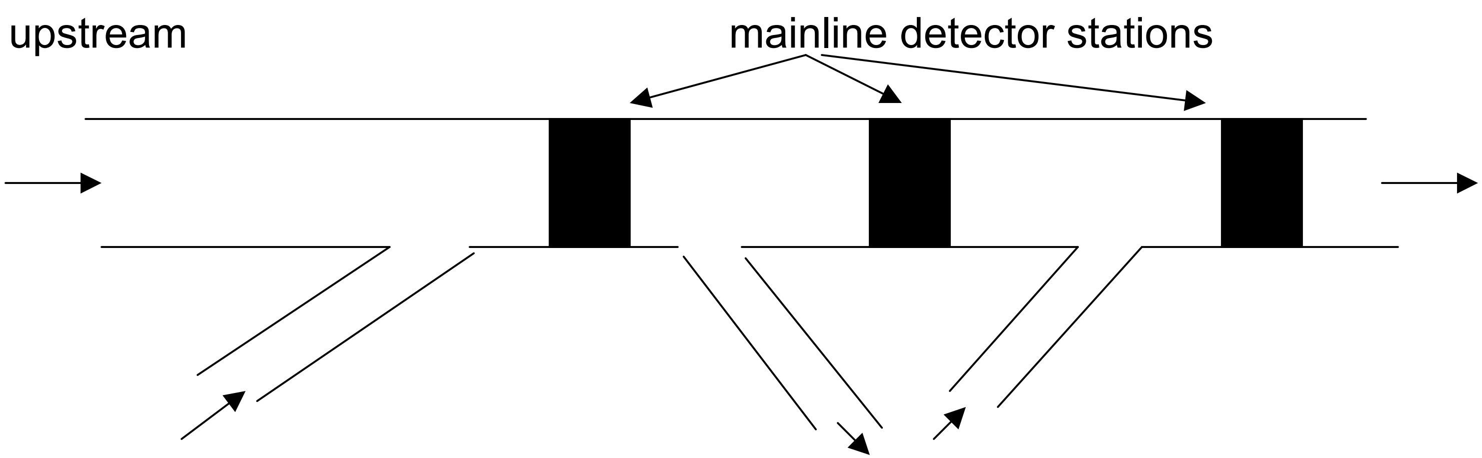
- VERY TIME CONSUMMING PROCESS
  - Currently Simulators Do Not Provide a Methodology or Tools to Assist in Calibration
  - Often Users End Up in Endless Trial-and-Error Cycles
  - Sporadic Attempts Made in Literature to Streamline the Process But:
    - Focused on a Particular Simulator
    - Too Complex or too Naive to be Effectively Used in Practice
    - No Widely Accepted Methods/Standards Currently Available

Before Calibration!

- Check Geometry For Correctness
  - Disjoined Sections
  - Stuck Vehicles (Sizes of Accel/Decel Lanes)
  - Verify Location of Detectors
- Check Input For Accuracy
  - Entrance Volume Comparison (Perfect Match)
  - Exit Volume Comparison (Match Sum Over All Hours)
  - Volume Totals on Mainline Stations Should Match

Practical Calibration Methodology (Employed on Twin Cities Freeways)

- Need Simultaneous Boundary and Mainline Station Measurements.
- Technique:



- Objective is to Match the Simulated and Actual Mainline Traffic Measurements
- Traffic Measurements Used: Volume and Speed
  - Occupancy Affected by Detector Sensitivity (Unknown)
- Perform Calibration in Stages:
  - First 2 Stages Based on Volume and Speed in That Order
  - Further Improvements in Optional 3rd Stage:
    - Depending on Objective i.e. For Ramp Control -> Queue Length

Goodness-of-Fit Test Measures

- Recommended Goodness-of-Fit Measures:

1. RMS Percent Error = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - y_i}{y_i}\right)^2}$
   (Measures Overall % Error)

2. Correlation Coefficient = $CORL = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}$
   (Measures Linear Association)

Where x_i is the Simulated Traffic Measurement at Time i
x' is the Mean of the Simulated Traffic Measurements
y_i is the Actual Traffic Measurement at Time i
y' is the Mean of the Actual Traffic Measurements
σ_x is the Standard Deviation of the Simulated Traffic Measurement
σ_y is the Standard Deviation of the Actual Traffic Measurements
n is the Number of Traffic Measurement Observations

Goodness-of-Fit Measures (Cont.)

3. Theil's U (Considers the Disproportionate Weight of Large Errors)

$$\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}y_i^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}}$$

- 3 Components of Theil's U
  - Us

    (Measure of Variance Proportion, Close to 1 Satisfactory)

$$U_s = \frac{n(\sigma_y - \sigma_x)^2}{\sum_{i=1}^{n}(y_i - x_i)^2}$$

  - Uc

    (Measure of Covariance Proportion or Unsystematic Error, Close to 0 Satisfactory)

$$U_c = \frac{2(1-r)n\sigma_y\sigma_x}{\sum_{i=1}^{n}(y_i - x_i)^2}$$

  - Um

    (Measure of Bias Proportion or Systematic Error, Close to 0 Satisfactory)

$$U_m = \frac{n(\overline{y} - \overline{x})^2}{\sum_{i=1}^{n}(y_i - x_i)^2}$$

Where σ_y and σ_x are the Standard Deviations of the Actual and Simulated Series

r is the Correlation Coefficient Between the Two Series

Examples

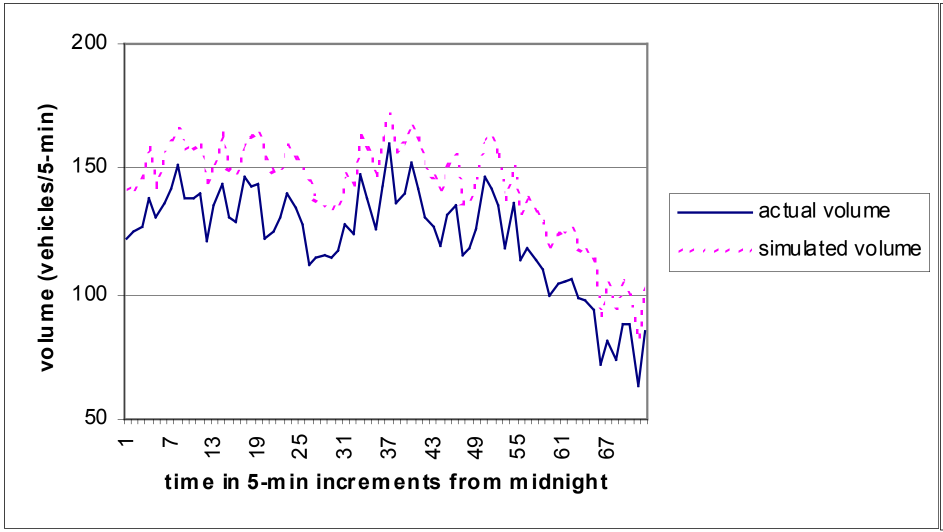


Figure 1(a): Illustration of unsatisfactory Um

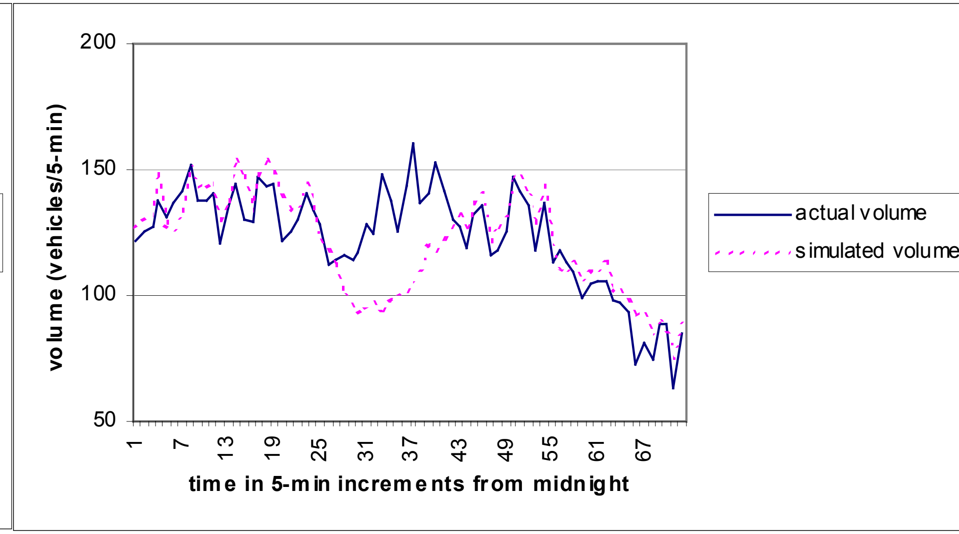


Figure 1(b): Illustration of unsatisfactory Uc

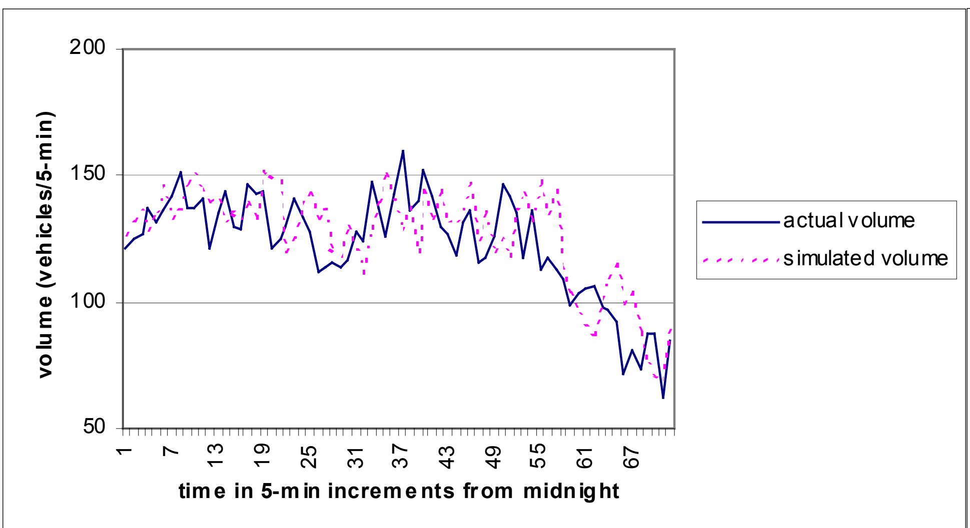


Figure 1(c): Illustration of unsatisfactory Uc

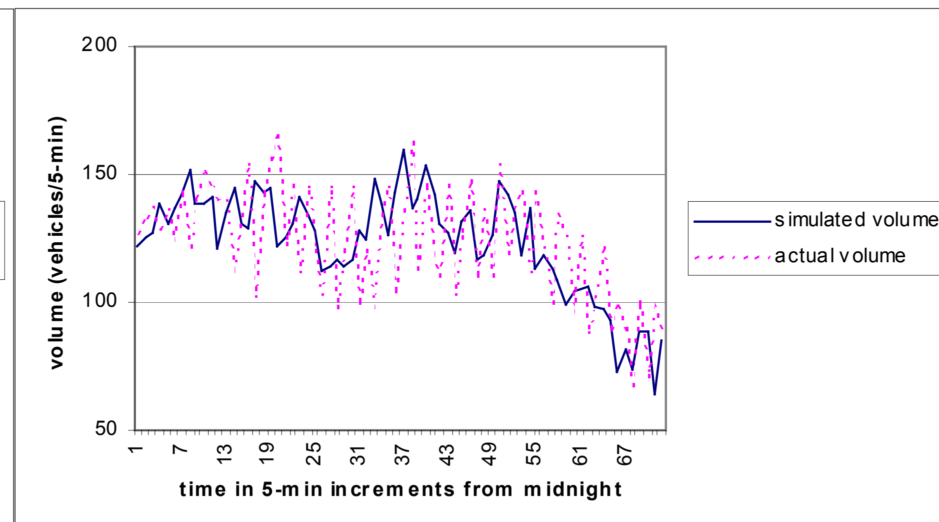


Figure 1(d): Illustration of unsatisfactory Us

Stage 1: Volume-Based Calibration

- Objective is to Match Simulated and Actual Mainline Station Volumes
- Simulation Model Calibrated Beginning Upstream and Proceeding Downstream
- Global Parameters Are Calibrated First:
  - Usually Accomplished in First Few Stations
  - Trial & Error Iterative Process For Each Parameter
  - RMSP, r and U are the Metrics Used in Each Iterations
- Local Parameters Calibrated at All Stations
  - Um, Uc and Us are the Metrics Used at This Point.

Stage 2: Speed-based Calibration

- Objective is to Match Simulated and Actual Mainline Station Speeds <u>and</u> Bottleneck Locations
- Actual Speeds Derived From Volume, Occupancy, and Effective Vehicle Length (For Single Loop Detectors)
- Speed Contour Graphs Used For Comparing (Visually) Simulated and Actual Speeds
  1. If Speed Contours Exhibit Significant Discrepancy, Revise Global Parameters From Stage 1
  2. Beginning Upstream and Proceeding Downstream, Calibrate Local Parameters Until Mainline Speeds and All Bottleneck Locations in the 2 Contour Graphs Match

Stage 3: Application Dependent

- For Adaptive Ramp control:
  - Compare Queue Lengths
  - Local Simulation Parameters Affecting Detector Output Need Further Calibration in This Stage

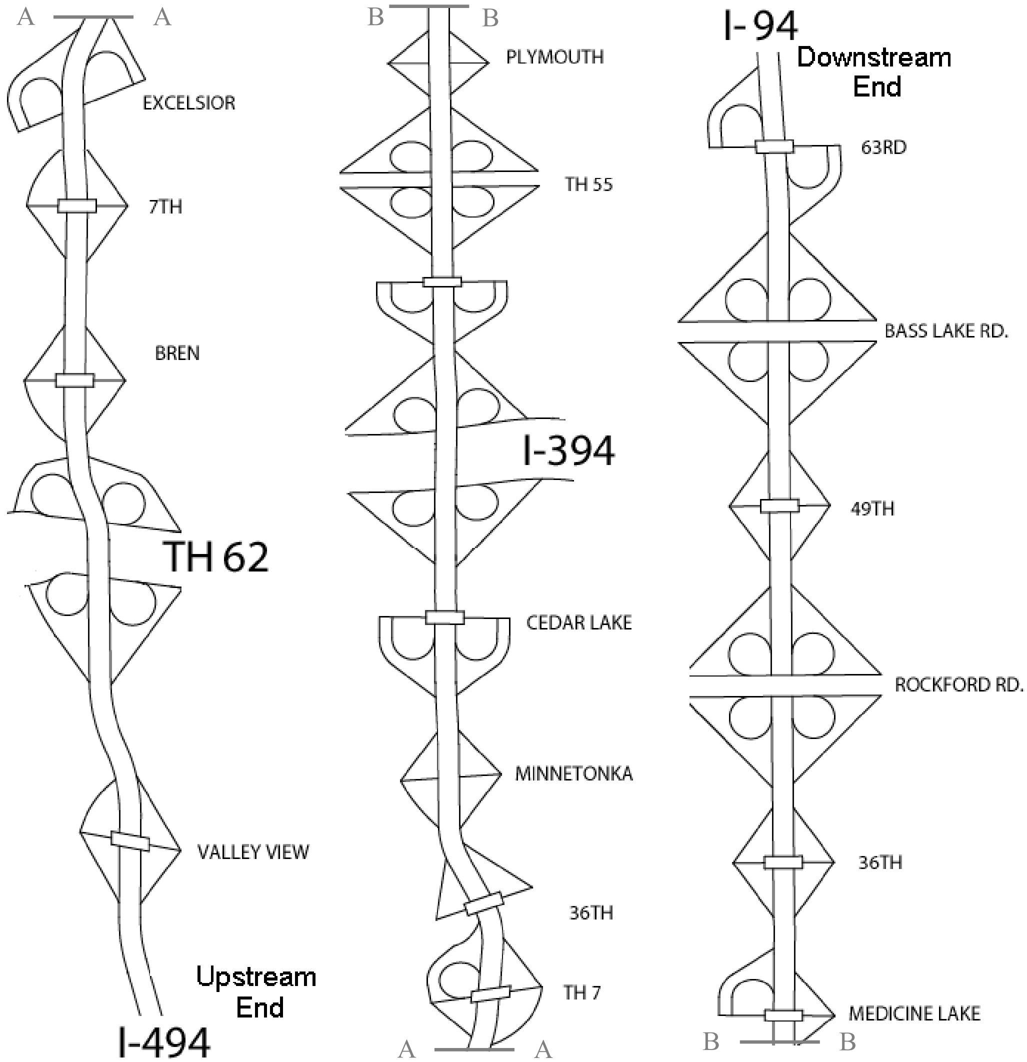
- Simulated Entrance Ramp Queues Should Match Actual Ones
- Queue Measurements From 30 sec Detector Counts
- Queue = Metered Demand - Actual Demand (Upstream)

- Over-Calibration to be Avoided to Ensure Generality
  - Repeat Simulation With Different Random Seeds
  - Simulate Additional Days

Example of Calibration

- AIMSUN Microsimulator
- Mn/DOT Ramp Metering Evaluation.
- 2 Test Sites (Only One Presented)
  - TH 169 Northbound in Minneapolis, MN
    - 12 Miles Long: From I-494 to 63rd Avenue N
    - 24 Entrance Ramps, 25 Exit Ramps
    - 30 Detector Stations.
    - 5-Minute Volume and Occupancy
    - March 21st to 23rd, 2000
    - 14:00 to 20:00 hrs

Test site 1: TH 169NB

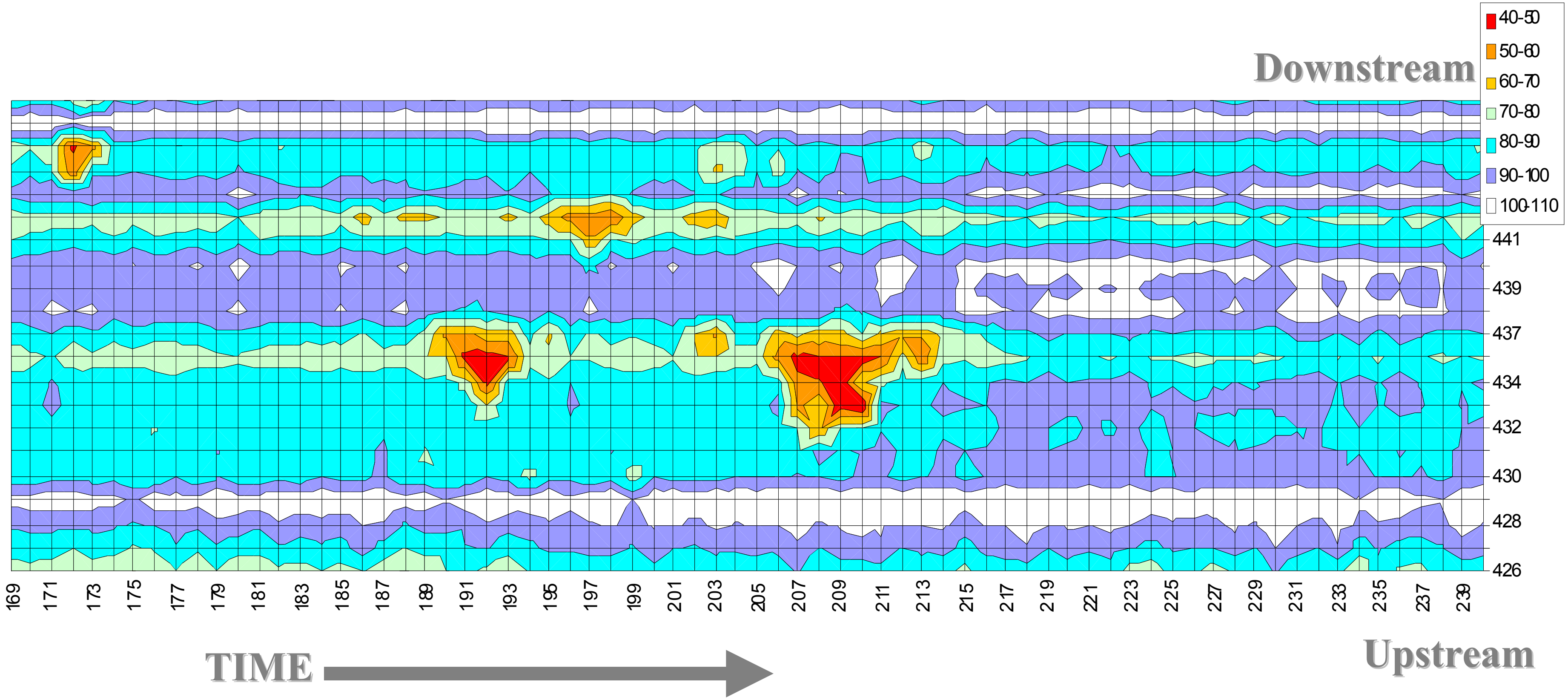


Stage 1 Results (Volume)

Goodness-of-Fit Measure	Values
r	0.960
RMSP	7.39%
U	0.002
Um	0.088
Us	0.031
Uc	0.881

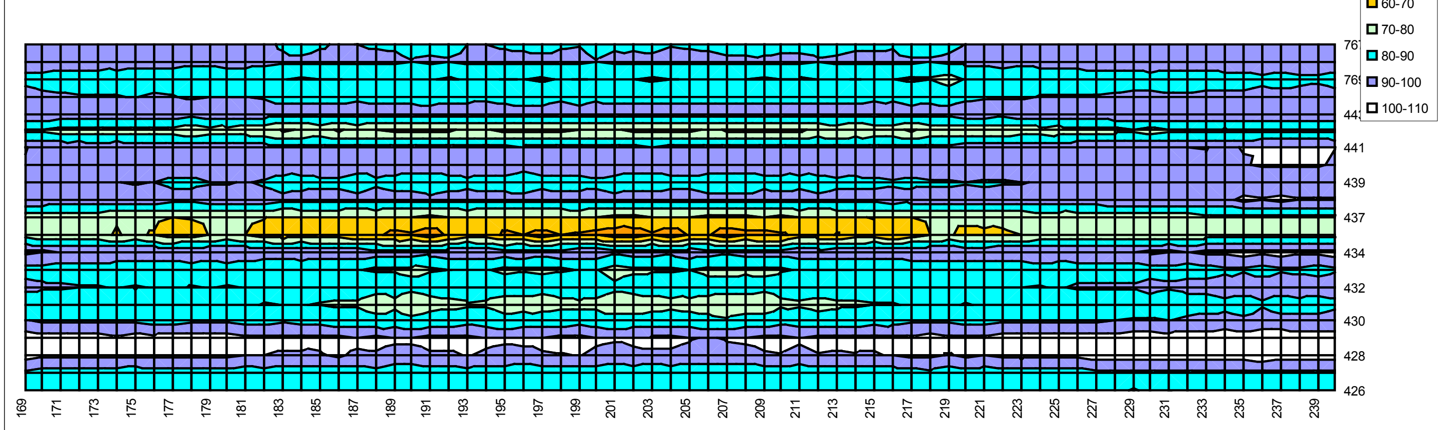
- 500 Simulator Iterations Required, 2 Months (!)
- Irregularities in Input Data Observed Due to Sensor Misplacement

Actual Speed Contour

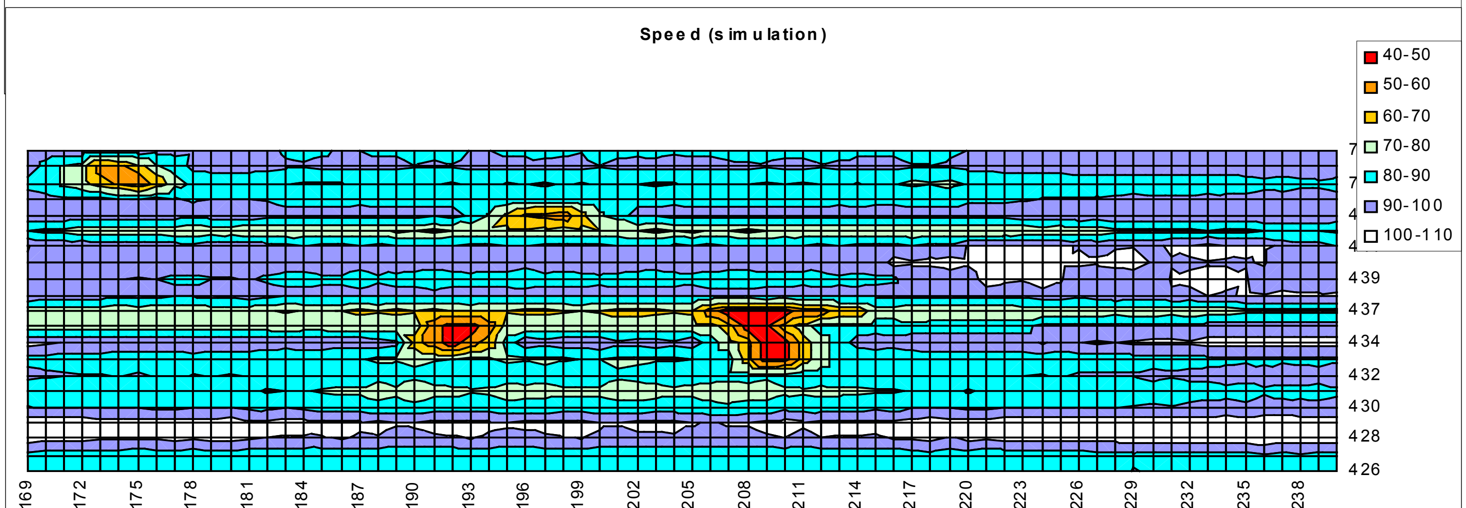


Stage 2 Results (Speed)

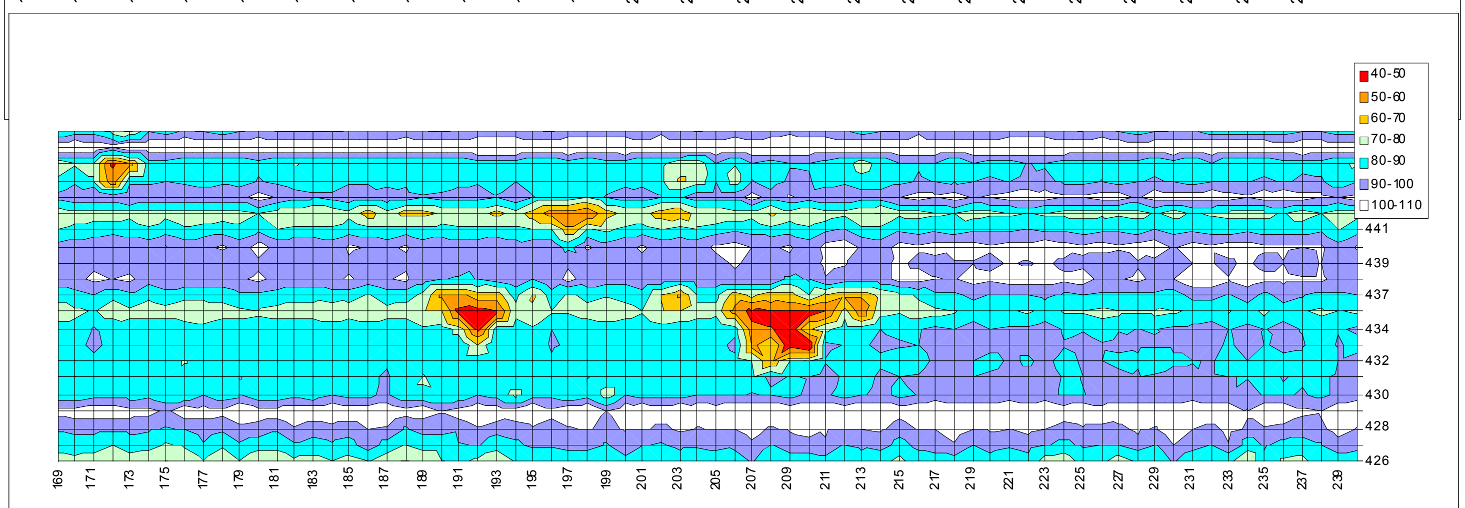
Simulated Speed After Stage 1



Simulated Speed After Stage 2 (About 200 Iterations)



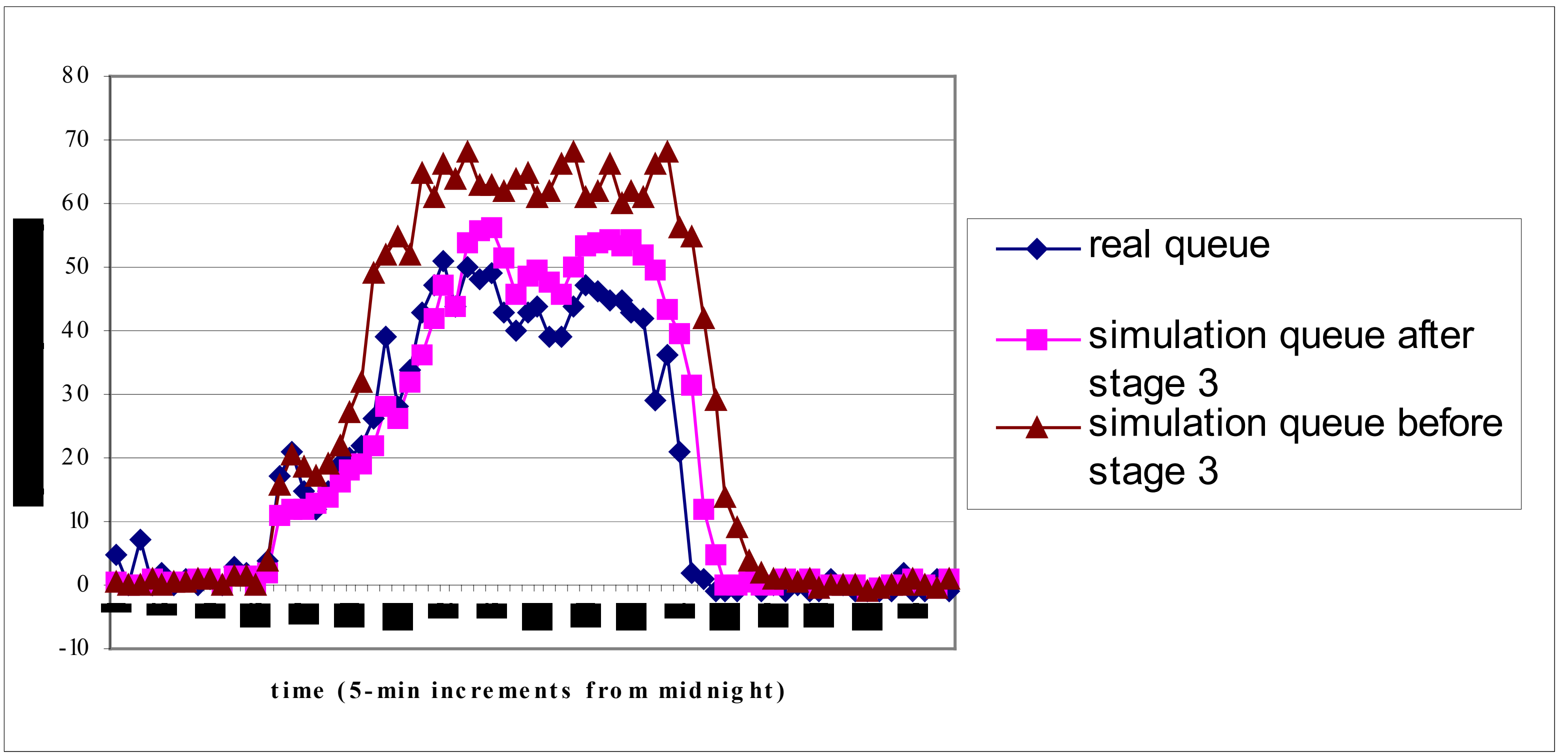
Actual Speed



Stage 3 Results (Queues)

- Simulated and Actual Queues Did Not Match Before This Stage
- 3rd Stage Required About 100 Iterations

Example Ramp



Simulation vs Real Queues

Validation Accuracy (Volume)

	RMS%	r	U	Um	Us	Uc
March 21st (validation)	10.620	0.980	0.004	0.309	0.011	0.681
March 22nd (validation)	6.420	0.970	0.001	0.124	0.054	0.823
March 23rd (Calibration)	7.390	0.960	0.002	0.088	0.031	0.881

Over All Stations.

Results (Calibrated Parameters) AIMSUN Microscopic Simulator

Parameter	Initial	After stage 1	After stage 2	After stage 3
Max. desired speed (kmph)	105.000	110.000	110.000	110.000
Max. acc. rate (m/s^2)	4.500	3.000	3.000	3.000
Normal dec. rate (m/s^2)	-4.500	-5.000	-5.000	-5.000
Max. dec. rate (m/s^2)	-5.000	-5.500	-5.500	-5.500
Reaction time (sec)	0.700	0.590	0.610	0.610
Percent overtake	0.950	0.950	0.940	0.940
Percent recover	1.000	1.000	0.990	0.990
Max. speed difference (kmph)	40.000	40.000	60.000	60.000
Max. speed difference on-ramp (kmph)	50.000	50.000	70.000	70.000
Av. section speed (regular section, kmph)	110.000	100.000	105.000	105.000
Av. section speed (weaving section, kmph)	90.000	75.000	70.000	70.000
Av. section speed (ramp section, kmph)	60.000	60.000	55.000	55.000

Conclusion

- Garbage In >>> Garbage Out
- Simulation Useless/Dangerous Without Calibration

Questions?